

Exploiting Delay Correlation for Multi-Antenna-Assisted High Speed Train Communications

Chunxu Jiao, Zhaoyang Zhang[†], Huazi Zhang, Liangliang Zhu

Department of Information Science and Electronic Engineering, Zhejiang University, Hangzhou 310027, China

E-mail: {jiaocx1990, ning_ming, hzhang17, zllzju}@zju.edu.cn

Abstract—In High Speed Train Communications (HSTC), the most challenging issue is coping with the extremely fast fading channel. Compared with its static counterpart, channel estimation on the move consumes excessive energy and spectrum to achieve similar performance. To address this issue, we exploit the *delay correlation* inherent in the linear spatial-temporal structure of multi-antenna array, based on which the rapid fading channel may be approximated by a virtual slow-fading channel. Subsequently, *error probability* and *spectral efficiency* are re-examined for this staticized channel. In particular, we formulate the quantitative tradeoff between the two metrics of interest, by adjusting the *pilot percentage* in each frame. Numerical results verify the good performance of the proposed scheme and elucidate the tradeoff.

I. INTRODUCTION

Recently, there is an emerging research trend towards communication scenarios with high node mobility, such as high speed train and highway communications. To achieve high data-rate in these scenarios, we face multiple challenges, e.g., fast handover, time-varying channel modeling, doubly selective fading modeling, pilot design and channel estimation, etc. However, most of them arise from the ultra-fast channel fading caused by high mobility.

Huge efforts have been made to enhancing wireless communication performance in high mobility scenarios (see [1]–[13] and the references therein). Among them, [1], [2] focus on the mobile channel modeling, based on mathematical analysis and measurement data, respectively. [3]–[6] research into the design of pilot symbols in various mobile systems. Moreover, [7], being more innovative, takes on node mobility as a type of multiuser diversity, indicating that mobility brings opportunities as well as challenges. However, the intrinsic effect of mobility on communication still needs further study. In mobile communication, will the space-time correlation of the channels be more helpful or more harmful? How can we further mitigate the negative influence caused by the rapid fading channel? How may the MIMO structure affect the communication performance? These problems are of great importance and need for more investigations.

On the one hand, correlation model plays a fundamental element in mobile channel analysis [1], [14]–[16]. The space-time modeling framework represents the partial correlation between the channels as well as fast fading and time dispersion. Among the previous works, [1] models the Rician fading

channel as the combination of diffuse and line-of-sight (LOS) components. Under linear node mobility, there is an inspiring observation that distinct antennas at the same location will share identical CSI. However, the static scattering environment assumption in [1] no longer fits in the dynamic scattering considered in our HSTC setting. Therefore, a new spatial-temporal correlation model capturing the dynamic of ambient objects is required for further analysis.

On the other hand, channel estimation using pilot symbols is a fundamental approach for providing robust communication over time-varying channels. However, ultra-high node mobility imposes formidable difficulties in the accurate tracking of channel state information (CSI). A straightforward solution is increasing pilot percentage, which will result in less sub-channels for data transmission and therefore sacrificing the spectral efficiency. In [3] and a recent relevant work [6], pilot is designed to minimize the error probability or maximize the spectral efficiency. Nevertheless, further understanding, especially the quantitative tradeoff between error probability and spectral efficiency, has not been fully revealed in literature.

In this paper, we mainly start with the *delay correlation* phenomenon and an improved channel estimation strategy that takes advantage of the former, and then analyze the tradeoff relations among various metrics of interest (e.g., error probability, spectral efficiency and pilot percentage). Consider the downlink of HSTC that utilizes MIMO techniques and employs pilot-assisted channel estimation. The receive array is usually mounted on the outside of the train and is equipped with multiple antennas. In particular, the receive antennas form a linear array with uniform spacing. Note that the link from base station to the train is an extremely rapid fading channel, which poses a major challenge to high mobility communication. Nevertheless, by exploiting the proposed *delay correlation*, we convert the fast time-varying channels into an approximately static channel. While we focus on the scenario of High Speed Train Communications (HSTC), the results in this paper may apply to general communications with linear node mobility and linear antenna array.

The contributions of this paper are summarized as follows:

- The phenomenon of *delay correlation* is discovered. It captures the spatial-temporal channel correlation between antennas moving to the same position at different time. We formalize this *delay correlation* to facilitate robust

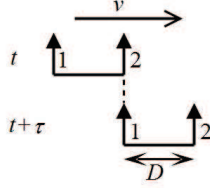


Fig. 1. A toy example of delay correlation: antenna 1 moves to the same position of antenna 2 after time τ , thus their channels to the source exhibit high correlation.

mobile communications, especially in HSTCs equipped with linear antenna array.

- We propose transmission schemes that exploit delay correlation to convert multiple fast fading channels to a virtual “static” channel, thus reducing the reliance on excessive pilot insertion. An analytical model is elaborated for this newly established “staticized” channel.
- Based on the “staticized” channel model, we analyze the tradeoff between error probability and spectral efficiency. This analytical tradeoff provides references for the parameter design in practical systems.

The rest of the paper is organized as follows. In Section II, the system model as well as the concept of *delay correlation* is presented. In section III, the impact of pilot percentage on error probability and spectral efficiency is analyzed, after that the tradeoff relation between error probability and spectral efficiency is elucidated. Section IV gives the relevant numerical results. Finally Section V concludes the work.

Throughout the rest of the paper, let $\mathbf{V} \in \mathbb{C}^{M \times N}$ means that the complex matrix \mathbf{V} is consist of M rows and N columns, the capital bold style means it is a matrix and the lowercase bold style means it is a vector. $\mathbb{E}[\cdot]$ is the mathematical expectation operator. $(\cdot)^*$, $(\cdot)^T$, $(\cdot)^H$ stand for complex conjugate, transpose and Hermitian transpose, respectively. $\|\cdot\|$ refers to the Frobenius norm. $\text{trace}(\cdot)$ is the matrix trace operation.

II. SYSTEM MODEL

In this section, the concept of *delay correlation* is proposed, based on which the fast fading channel can be approximated by an quasi-static channel. Remarkably, the frames are no longer composed of successive received symbols, but are formed with symbols that benefit from *delay correlation*.

A. Delay Correlation Model

Delay correlation is a unique phenomenon caused by linear mobility and linearly spaced antenna array. To be specific, high speed train has deterministic moving direction and relatively steady speed in a short time. In addition, the receive array, which may utilize massive MIMO techniques, can be positioned to form a line topology in practical design. The aforementioned two assumptions form the foundation for this entire work.

Consider the toy example in Fig. 1, where two receive antennas move along the line defined by themselves. Then when Antenna 1 moves to the current position of Antenna 2

after time τ , it experiences a similar channel that Antenna 2 experienced τ time ago¹. In high speed scenario, τ can be so short that above mentioned two channels are almost identical. So in this case if Antenna 2 informs Antenna 1 the channel it estimated τ time ago, then Antenna 1 may save the effort for additional channel estimation. This simple but inspiring phenomenon offers even bigger advantage when the antenna number increases, as we will analyze in detail later.

Regarding correlation modeling, static scattering environment is assumed in previous works, which indicates that the channel will remain unchanged as long as the concerned antennas are immobile. However, if the ambient objects are dynamic, intuitively, there will be a decay factor attached to the cross-correlation coefficient of the diffuse counterpart. In particular, provided with the delay τ and antenna spacing D for the above example, the correlation coefficient of Antenna 2 at time t and Antenna 1 at time $t + \tau$ is given by the following definition.

Definition: The cross-correlation function for mobile frequency non-selective 1×2 Rician fading MIMO channels is

$$\rho^{\text{DIF}}(\tau, D) = \frac{I_0 \left(\sqrt{\kappa^2 - \Delta^2} - j2\kappa\Delta \cos(\mu - \gamma) \right)}{(K_R + 1) I_0(\kappa)} e^{-c_0 v |\tau|}, \quad (1)$$

$$\rho^{\text{LOS}}(\tau, D) = \frac{K_R}{K_R + 1} e^{j\Delta \cos(\gamma)}, \quad (2)$$

where $\Delta = 2\pi \left(f_D \tau - \frac{D}{\lambda} \right)$ presents the location difference of the two antennas, $f_D = \frac{v}{\lambda}$ is the maximum Doppler shift and λ is the signal wavelength. κ controls the width of angle of arrival (AOA) and $\mu \in [-\pi, \pi)$ accounts for the mean direction of AOA. The mobility of the train is characterized by its velocity v and direction γ . Finally, c_0 is a real-valued constant characterizing the inherent spatial property of the scattering objects.

Remarks: Some explanations for this concept are in order.

1) The cross-correlation of the diffuse part is mainly characterized by three factors: a) location difference; b) AOA difference; c) ambient objects' movement. Apparently, static antennas and static???? scattering environment result in perfect correlation.

2) The LOS part only varies in phase of arrival, and the phase shift is obtained easily through geometric analysis.

3) This definition is in accordance with the result in [1] and may naturally be extended to multiple, even massive, antennas topology.

Based on the above definition, if τ is set as $\frac{D}{v}$, then Δ will be 0, and finally, *delay correlation* is formulated. To be specific, *delay correlation* is characterized by

$$\rho^{\text{DIF}}(\tau, D) \big|_{\tau=\frac{D}{v}} = \frac{1}{K_R + 1} e^{-c_0 D} \quad (3)$$

$$\rho^{\text{LOS}}(\tau, D) \big|_{\tau=\frac{D}{v}} = \frac{K_R}{K_R + 1} \quad (4)$$

Note that the direction and velocity of the train are known by both of the base station and the train.

¹For simplicity, assume that source S is equipped with single antenna.

where $\hat{\mathbf{h}}_p^{\text{DIF}} \in \mathbb{C}^{N_p \times 1}$ is the estimation of $\mathbf{h}_p^{\text{DIF}}$, and $\mathbf{W}_p = \sqrt{E_0}(E_0 \mathbf{P} \mathbf{R}_{hh}^{\text{DIF}} \mathbf{P}^H + \sigma_n^2 \mathbf{I}_{N_p})^{-1} \mathbf{P} \mathbf{R}_{hh}^{\text{DIF}}$ is the channel MMSE estimation matrix, auto-correlation matrix of the diffuse channel $\mathbf{R}_{hh}^{\text{DIF}} = \mathbb{E}[\mathbf{h}_p^{\text{DIF}} (\mathbf{h}_p^{\text{DIF}})^H] \in \mathbb{R}^{N_p \times N_p}$ is a Toeplitz matrix whose (m, n) -th element being

$$\rho_{m,n} = \rho_{m,n}^{\text{DIF}} = \frac{1}{K_R + 1} e^{-c_0 |m-n| \frac{\delta}{\delta}}, \quad (12)$$

where δ is the pilot percentage.

In the second step, the channel estimation of arbitrary h_i is obtained through interpolating $\hat{\mathbf{h}}_p^{\text{DIF}}$ and attaching the LOS component

$$\hat{h}_i = \mathbf{w}_d^H \hat{\mathbf{h}}_p^{\text{DIF}} + h_i^{\text{LOS}}, \quad (13)$$

where the real-valued coefficient vector $\mathbf{w}_d \in \mathbb{R}^{N_p \times 1}$ is designed according to the MSE minimization criterion (9).

It is proved in [17] that this scheme is equivalent to the optimum linear MMSE estimator as follows

$$\hat{h}_i = \sqrt{E_0} \mathbf{r}_i^{\text{DIF}} \mathbf{P}^H (E_0 \mathbf{P} \mathbf{R}_{hh}^{\text{DIF}} \mathbf{P}^H + \sigma_n^2 \mathbf{I}_{N_p})^{-1} \times (\mathbf{y}_p - \sqrt{E_0} \mathbf{P} \mathbf{h}_p^{\text{LOS}}) + h_i^{\text{LOS}}, \quad i = 1, \dots, N_R, \quad (14)$$

where $\mathbf{r}_i^{\text{DIF}} = \mathbb{E}[\mathbf{h}_i^{\text{DIF}} (\mathbf{h}_p^{\text{DIF}})^H]$, which indicates that $\mathbf{w}_d^H = \mathbf{r}_i^{\text{DIF}} (\mathbf{R}_{hh}^{\text{DIF}})^{-1}$.

B. Error Probability vs. Pilot Percentage

The accuracy of channel estimation is mainly determined by pilot percentage. Meanwhile, imperfect channel estimation results in high error probability [18]. Hence, there exists a mapping between error probability P_e and pilot percentage δ .

Above all, the MSE of the *Two-Step MMSE* estimation scheme is derived. As stated in [6], define the error correlation matrix at pilot locations as

$$\mathbf{R}_{ee} = \mathbb{E}[\mathbf{e}_p \mathbf{e}_p^H], \quad (15)$$

where $\mathbf{e}_p = \hat{\mathbf{h}}_p - \mathbf{h}_p = \hat{\mathbf{h}}_p^{\text{DIF}} - \mathbf{h}_p^{\text{DIF}}$. For simplicity, assume that the pilot symbols $\mathbf{P} = \mathbf{I}_{N_p}$. Substituting this assumption and (10)(11) into (15), it can be calculated that

$$\mathbf{R}_{ee} = \mathbf{R}_{hh}^{\text{DIF}} - \mathbf{R}_{hh}^{\text{DIF}} \left(\mathbf{R}_{hh}^{\text{DIF}} + \frac{1}{\gamma} \mathbf{I}_{N_p} \right)^{-1} \mathbf{R}_{hh}^{\text{DIF}}, \quad (16)$$

where $\gamma = \frac{E_0}{\sigma_n^2}$ is the signal-to-noise ratio (SNR). Then the average MSE at pilot locations is

$$\sigma_{p,N_p}^2 = \frac{1}{N_p} \text{trace}(\mathbf{R}_{ee}), \quad (17)$$

Through asymptotic analysis, i.e., $N_p \rightarrow \infty$, $N_s \rightarrow \infty$ and $N_R \rightarrow \infty$, with a finite pilot percentage δ , the asymptotic MSE at pilot locations can be expressed as a function of δ with the following theorem. Note that asymptotic analysis is reasonable for massive MIMO where a large amount of receiver antennas exist.

Theorem 1. *Let $N_p \rightarrow \infty$ and $N_s \rightarrow \infty$ while keeping a finite pilot percentage δ , the asymptotic MSE $\sigma_p^2 = \lim_{N_p \rightarrow \infty} \sigma_{p,N_p}$ at pilot locations is*

$$\sigma_p^2 = \sqrt{\frac{1}{\gamma^2 + 2\gamma \frac{1+\alpha^2}{1-\alpha^2} (K_R + 1) + (K_R + 1)^2}}, \quad (18)$$

where $\alpha = e^{-c_0 \frac{\delta}{\delta}}$.

Proof: The proof is in Appendix A. ■

Remarks: We can see that the estimation MSE will decrease when SNR and Rician factor rise. And the increase in pilot percentage leads to higher correlation between channels at adjacent pilot locations. Moreover, stronger intraclass correlation facilitates more accurate estimations. Therefore, increase in pilot percentage will concurrently result in smaller MSE. These observations will be further verified through the simulation results.

Next, the MSE of arbitrary h_i^{DIF} needs to be calculated. Decompose the remaining N_s data symbols into L groups, each with N_p symbols.

Define the u -th symbol group as $\mathbf{X}_u \in \mathbb{C}^{N_p \times N_p}$, which contains the data symbols with indices $\{i'_k\} = (k-1)(L+1) + u$, $k = 1, \dots, N_p$. In addition, define its corresponding channels and received symbols as $\mathbf{h}_{d,u}$ and $\mathbf{y}_{d,u}$, respectively. In this case, the system model becomes

$$\mathbf{y}_{d,u} = \sqrt{E_0} \mathbf{X}_u \mathbf{h}_{d,u} + \mathbf{n}_u, \quad (19)$$

From (13), the MMSE estimation of $\mathbf{h}_{d,u}$ is

$$\hat{\mathbf{h}}_{d,u} = \mathbf{W}_{d,u}^H \hat{\mathbf{h}}_p^{\text{DIF}} + \mathbf{h}_{d,u}^{\text{LOS}}, \quad (20)$$

where $\mathbf{W}_{d,u}^H = \mathbf{R}_{dh,u}^{\text{DIF}} (\mathbf{R}_{hh}^{\text{DIF}})^{-1}$ contains the interpolating coefficients, and $\mathbf{R}_{dh,u}^{\text{DIF}} = \mathbb{E}[\mathbf{h}_{d,u}^{\text{DIF}} (\mathbf{h}_p^{\text{DIF}})^H] \in \mathbb{R}^{N_p \times N_p}$ is also a Toeplitz matrix whose (m, n) -th element is

$$\rho_{m,n} = \rho_{m,n}^{\text{DIF}} = \frac{1}{K_R + 1} e^{-c_0 |m-n+u\delta| \frac{\delta}{\delta}}, \quad (21)$$

Then the corresponding error correlation matrix of channel estimations at these symbol locations, defined as $\Psi_{ee,u} = \mathbb{E}[\mathbf{e}_{d,u} \mathbf{e}_{d,u}^H]$ with $\mathbf{e}_{d,u} = \hat{\mathbf{h}}_{d,u} - \mathbf{h}_{d,u} = \hat{\mathbf{h}}_{d,u}^{\text{DIF}} - \mathbf{h}_{d,u}^{\text{DIF}}$, can be calculated as

$$\Psi_{ee,u} = \mathbf{R}_{dd,u}^{\text{DIF}} - \mathbf{R}_{dh,u}^{\text{DIF}} \left(\mathbf{R}_{hh}^{\text{DIF}} + \frac{1}{\gamma} \mathbf{I}_{N_p} \right)^{-1} (\mathbf{R}_{dh,u}^{\text{DIF}})^H, \quad (22)$$

where $\mathbf{R}_{dd,u}^{\text{DIF}} = \mathbb{E}[\mathbf{h}_{d,u}^{\text{DIF}} (\mathbf{h}_{d,u}^{\text{DIF}})^H] = \mathbf{R}_{hh}^{\text{DIF}}$. Similarly, the average MSE for the channel estimations at u -th group data symbol locations is

$$\sigma_{d,u,N_p}^2 = \frac{1}{N_p} \text{trace}(\Psi_{ee,u}), \quad (23)$$

Through asymptotic analysis, the MSE at data symbol locations is given by the following theorem.

Theorem 2. *Let $N_p \rightarrow \infty$ and $N_s \rightarrow \infty$ while keeping a finite pilot percentage δ , the asymptotic MSE $\sigma_{d,u}^2 = \lim_{N_p \rightarrow \infty} \sigma_{d,u,N_p}$ at u -th group data symbol locations is*

$$\sigma_{d,u}^2 = \sigma_p^2 + \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\Lambda(\Omega)^2 - |\Lambda_{dh,u}(\Omega)|^2}{\Lambda(\Omega) + \frac{1}{\gamma}} \right] d\Omega, \quad (24)$$

where $\Lambda(\Omega)$ and $\Lambda_{dh,u}(\Omega)$ are as follows

$$\Lambda(\Omega) = \frac{1}{K_R + 1} \left[\frac{1 - \alpha^2}{1 - 2\alpha \cos(\Omega) + \alpha^2} \right], \quad (25)$$

$$\Lambda_{dh,u}(\Omega) = \frac{1}{K_R + 1} \left[\frac{\alpha(\beta^{-1} - \beta) e^{j\Omega} + \beta - \alpha^2 \beta^{-1}}{1 - 2\alpha \cos(\Omega) + \alpha^2} \right], \quad (26)$$

$$\alpha = e^{-c_0 \frac{D}{8}} \text{ and } \beta = e^{-c_0 u D}.$$

Proof: The proof is in Appendix B. ■

Remarks: In low SNR regime, the estimation error will be mainly dominated by noise, while β , the correlation coefficient between the channels of the u -th symbol group and the prior pilot symbols, has negligible effect on the MSE. In this case, the second item of the expression would be very close to 0, indicating that the MSE at data locations is quite similar to that at pilot locations. However, in high SNR regime, channels at pilot locations are perfectly estimated, so noise will no longer affect the estimation error while *delay correlation* becomes more predominating. In this sense, the first item approaches 0 and the second item becomes almost a positive constant, along which the diversity order would be zero.

Having obtained the analytical MSEs, it is possible to analyze the error probability of our system model. For BPSK, error probability P_e is equal to BER. The maximum likelihood decision rule is

$$\hat{x}_i = \arg \min_{x \in \{-1, 1\}} \left\{ \left| \frac{1}{\sqrt{E_0}} h_i^* y_i - x \right| \right\}, \quad (27)$$

Thus the BER performance is given in [19]

$$P_{e,u} = \frac{1}{\pi} e^{-\frac{K_R}{\rho^2}} \int_0^{\frac{\pi}{2}} \left[1 + \frac{\tilde{\gamma}}{K_R + 1} \frac{1}{\sin^2(\phi)} \right]^{-1} \times \exp \left\{ \frac{K_R}{\rho^2} \left[1 + \frac{\tilde{\gamma}}{K_R + 1} \frac{1}{\sin^2(\phi)} \right]^{-1} \right\} d\phi, \quad (28)$$

where $\rho = \sqrt{\frac{1}{1 + (K_R + 1)\sigma_{d,u}^2}}$ is the covariance coefficient between $\mathbf{h}_{d,u}$ and $\hat{\mathbf{h}}_{d,u}$, $\tilde{\gamma} = \frac{(K_R + 1)\rho^2}{\gamma(1 - \rho^2) + K_R + 1} \gamma = \frac{1}{\sigma_{d,u}^2 + \frac{1}{\gamma}}$ is the equivalent SNR.

As $\sigma_{d,u}^2$ is related to the index u , the average BER is simply derived by calculating the mean of $P_{e,u}$ over u

$$P_e = \mathbb{E}_u [P_{e,u}] = \frac{1}{L} \sum_{u=1}^L P_{e,u}. \quad (29)$$

C. Spectral Efficiency vs. Pilot Percentage

Intuitively, more pilot symbols result in better channel estimations, which would help improve the spectral efficiency. On the other hand, excessive pilot symbols would lead to unnecessary spectrum overhead. Hence, there exists an optimal pilot percentage which would maximize the spectral efficiency. [6] gives the maximum spectral efficiency analysis with imperfect channel information in Rayleigh fading channels. In a similar fashion, the relation between spectral efficiency η and pilot percentage δ in Rician fading models will be obtained.

In this paper, the effective spectral efficiency is defined as

$$\eta = \frac{N_s}{N_R} \mathbb{E}_{\hat{h}} [C(\hat{h}, \gamma)] = (1 - \delta) \mathbb{E}_{\hat{h}} [C(\hat{h}, \gamma)], \quad (30)$$

where $C(\hat{h}, \gamma)$ is the channel capacity with imperfect channel estimations \hat{h} and SNR γ .

Lemma 1. For a M-PSK modulation system operating in a Rician fading channel with pilot-assisted MMSE channel estimation with regard to the u -th symbol group, the channel capacity conditioned on imperfect channel estimation \hat{h} and SNR γ is upper bounded by

$$C(\hat{h}, \gamma)_{u,up} = \log(1 + |\hat{h}|^2 \tilde{\gamma}), \quad (31)$$

where $\tilde{\gamma} = \frac{1}{\sigma_{d,u}^2 + \frac{1}{\gamma}}$ is the equivalent SNR.

Proof: The proof is in Appendix C. ■

Based on the above lemma, take the average of $C(\hat{h}, \gamma)_{u,up}$ over \hat{h} and u , and then substitute the expectation into (30), the expression of effective spectral efficiency is derived.

Theorem 3. For the system operating in a Rician fading channel with pilot-assisted MMSE channel estimation with regard to the u -th symbol group, the effective spectral efficiency is given by

$$\eta_{u,up} \approx (1 - \delta) \int_0^{+\infty} C(x) p(x) dx, \quad (32)$$

where $C(x)$ and $p(x)$ is expressed as

$$C(x) = \log \left(1 + \frac{x}{2(K_R + 1)} \tilde{\gamma} \right), \quad (33)$$

$$p(x) = \frac{1}{2} e^{-\frac{x+2K_R}{2}} I_0(\sqrt{2K_R x}), \quad (34)$$

and $\tilde{\gamma} = \frac{1}{\sigma_{d,u}^2 + \frac{1}{\gamma}}$ is the equivalent SNR.

Proof: The proof is in Appendix D. ■

Remarks: Apparently, the spectrum is used more efficiently in high SNR regime. Furthermore, as the pilot percentage increases, data percentage will keep decreasing while the corresponding capacity of the data channel maintains rising. Note that the effective spectral efficiency is the product of data percentage and the average data channel capacity, therefore, there exists an optimal pilot percentage that maximizes the spectral efficiency.

Similarly, the spectral efficiency should be averaged on u , thus getting the final result

$$\eta_{up} = \frac{1}{L} \sum_{u=1}^L \eta_{u,up}. \quad (35)$$

D. Error Probability vs. Spectral Efficiency

After SNR γ and pilot percentage δ are specified, the corresponding MSE at data symbol locations can be calculated through (24). Furthermore, the average error probability P_e and spectral efficiency η_{up} can be derived through (29) and (35), respectively. Evidently, it is unrealistic to simultaneously achieve lower P_e and higher η_{up} , thus implying a tradeoff relation between them.

According to (32), it is possible to obtain the maximum spectral efficiency through selecting a specific δ_{SE-opt} , yet the corresponding BER performance would not necessarily be satisfactory. However, it is worth nothing that δ_{SE-opt} has practical implication in our tradeoff analysis, i.e., it would serve as the lower bound of feasible pilot percentages. When

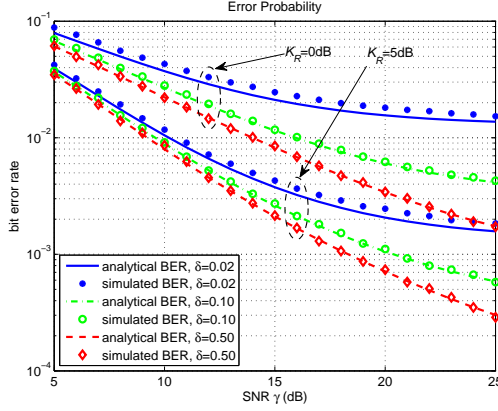


Fig. 3. error probability as a function of SNR γ and pilot percentage δ .

$\delta > \delta_{\text{SE-opt}}$, both spectral efficiency and error probability would decrease monotonically. Hence, it is crucial to choose a pilot percentage between $\delta_{\text{SE-opt}}$ and 0.5 to balance between error probability and spectral efficiency.

Having understood the relation among MSE, error probability and spectral efficiency, it makes sense to adjust δ between $\delta_{\text{SE-opt}}$ and 0.5 while calculating the corresponding P_e and η_{up} , thus obtaining the optimal tradeoff between error probability and spectral efficiency.

IV. NUMERICAL RESULTS

Numerical results are provided in this section. Firstly, the impacts of pilot percentage on error probability and spectral efficiency are verified. After that, the tradeoff relation between error probability and spectral efficiency is elicited.

Fig. 3 demonstrates the relation between P_e and γ under various simulation parameters, i.e., $\delta = 0.02, 0.10, 0.50$ and $K_R = 0\text{dB}, 5\text{dB}$. The scattering objects spatial parameter $c_0 = 0.1\text{m}^{-1}$. We can see that there exists a negative relationship between P_e and γ , K_R as well as δ , which verifies Theorem 1, Theorem 2 and the error probability expressions (28)(29). Notably, it can be observed that there is an error floor for this communication system, the reason is that the MSE at data locations is independent of γ in high SNR regime.

Next, the relation between η_{up} and δ is elucidated in Fig. 4, in which the spectral efficiency rises first, and then falls. In this sense, the spectral efficiency can be maximized. However, with the pilot percentage as $\delta_{\text{SE-opt}}$, the BER performance is less than satisfactory, especially in high SNR regime. It is practical to select a slightly higher δ which would balance between error probability and spectral efficiency.

The tradeoff between error probability and spectral efficiency is depicted in Fig. 5. The feasible pilot percentages are several discrete values in range $[\delta_{\text{SE-opt}}, 0.5]$, along which P_e and η_{up} are both maximized when $\delta = \delta_{\text{SE-opt}}$. For normalization, P_e and η_{up} are divided by $P_e(\delta_{\text{SE-opt}})$ and $\eta_{up}(\delta_{\text{SE-opt}})$, respectively. From this plot, it can be seen that the tradeoff between P_e and η_{up} is elucidated as a polyline, the corresponding pilot percentages are $\frac{1}{2}, \frac{1}{3}, \dots, \delta_{\text{SE-opt}}$.

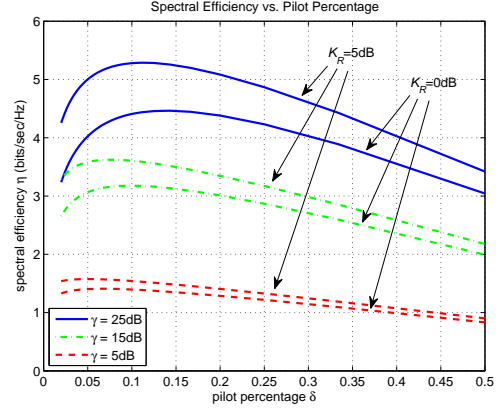


Fig. 4. The relation between spectral efficiency and pilot percentage.

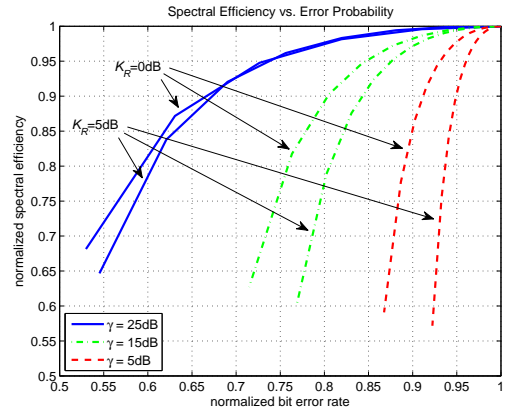


Fig. 5. The relation between spectral efficiency and error probability.

Moreover, when γ and K_R get high, a small loss of spectral efficiency will bring in conspicuous BER performance gain.

V. CONCLUSION AND FUTURE WORK

In this paper, *delay correlation* inherent in HSTC system is exploited to provide robust communication under high mobility while reducing the use of pilot symbols. A novel staticized channel model based on *delay correlation* is proposed and analyzed, which converts the rapid fading channel into a virtual static channel. In particular, we derived the closed-form expressions for two metrics of interest, i.e., error probability and spectral efficiency, and obtained an analytical tradeoff between them. This tradeoff provides useful references for the parameter design in future delay-correlation-enabled systems.

Delay correlation provide an alternative robust space-time communication solution to many existing systems. In our current scheme, it is assumed that only one antenna of the receive array is selected to be active each time. We plan to extend the number of activated antennas and obtain diversity gain.

APPENDIX

A. Proof of Theorem 1

(17) is equal to summarizing the eigenvalues of $\frac{1}{N_p} \mathbf{R}_{ee}$ as follows

$$\begin{aligned} \sigma_{p,N_p}^2 &= \frac{1}{N_p} \sum_{n=1}^{N_p} \left[\lambda_n - \left(\lambda_n + \frac{1}{\gamma} \right)^{-1} \lambda_n^2 \right] \\ &= \frac{1}{N_p} \sum_{n=1}^{N_p} \left(\frac{\lambda_n}{\lambda_n \gamma + 1} \right) \end{aligned} \quad (36)$$

where λ_n is the n -th eigenvalue of \mathbf{R}_{hh} . Based on the asymptotic analysis method in [20], when $N_p \rightarrow \infty$, σ_{p,N_p}^2 can be recast as

$$\sigma_p^2 = \lim_{N_p \rightarrow \infty} \sigma_{p,N_p}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\Lambda(\Omega)}{\Lambda(\Omega)\gamma + 1} \right] d\Omega \quad (37)$$

where $\Lambda(\Omega)$ is the discrete-time Fourier transform (DTFT) of $\left\{ \frac{1}{K_R+1} e^{-c_0|k|\frac{D}{\sigma}} \right\}_k$, which is expressed as

$$\begin{aligned} \Lambda(\Omega) &= \sum_{k=-\infty}^{\infty} \left(\frac{1}{K_R+1} e^{-c_0|k|\frac{D}{\sigma}} \right) e^{-jk\Omega} \\ &= \frac{1}{K_R+1} \left[\frac{1 - \alpha^2}{1 - 2\alpha \cos(\Omega) + \alpha^2} \right] \end{aligned} \quad (38)$$

where $\alpha = e^{-c_0\frac{D}{\sigma}}$. Therefore, σ_p^2 could be calculated by

$$\begin{aligned} \sigma_p^2 &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\frac{\frac{1}{K_R+1}}{\frac{\gamma}{K_R+1} + \frac{1+\alpha^2}{1-\alpha^2} - \frac{2\alpha}{1-\alpha^2} \cos(\Omega)} \right] d\Omega \\ &= \frac{1}{K_R+1} \times \sqrt{\frac{1}{a^2 - b^2}} \end{aligned} \quad (39)$$

where $a = \frac{\gamma}{K_R+1} + \frac{1+\alpha^2}{1-\alpha^2}$ and $b = -\frac{2\alpha}{1-\alpha^2}$.

B. Proof of Theorem 2

The same with (36) and (37), σ_{d,u,N_p}^2 can be rewritten as

$$\sigma_{d,u}^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left[\Lambda(\Omega) - \frac{|\Lambda_{dh,u}(\Omega)|^2}{\Lambda(\Omega) + \frac{1}{\gamma}} \right] d\Omega \quad (40)$$

$\Lambda_{dh,u}(\Omega)$ is the DTFT of $\left\{ \frac{1}{K_R+1} e^{-c_0|k+\delta u|\frac{D}{\sigma}} \right\}_k$ and is expressed as

$$\Lambda_{dh,u}(\Omega) = \frac{1}{K_R+1} \left[\frac{\alpha(\beta^{-1} - \beta) e^{j\Omega} + \beta - \alpha^2 \beta^{-1}}{1 - 2\alpha \cos(\Omega) + \alpha^2} \right] \quad (41)$$

where $\beta = e^{-c_0 u D}$. Substituting (38) and (41) into (40), $\sigma_{d,u}^2$ could be recast as the form in (24).

C. Proof of Lemma 1

The capacity $C(\hat{h}, \gamma)$ is equivalent to the maximum conditional mutual information $\max\{I(x; y|\hat{h})\}$, where $I(x; y|\hat{h})$ is defined below

$$I(x; y|\hat{h}) = E_{x,y} [\log p(y|x, \hat{h})] - E_y [\log p(y|\hat{h})] \quad (42)$$

where x is a symbol of the u -th group, \hat{h} and y are the corresponding estimated channel coefficient and received signal.

Conditional on \hat{h} and x , y is Gaussian distributed with mean and variance given in [6]

$$\mu_{y|x,\hat{h}} = \sqrt{E_0} \hat{h} x \quad (43)$$

$$\sigma_{y|x,\hat{h}}^2 = E_0 \sigma_{d,u}^2 + \sigma_n^2 \quad (44)$$

$|x_u| = 1$ is utilized in the second equation. Therefore,

$$E_{x,y} [\log p(y|x, \hat{h})] = \log \frac{1}{\pi e (E_0 \sigma_{d,u}^2 + \sigma_n^2)} \quad (45)$$

Note that $h - \hat{h} \sim \text{CN}(0, \sigma_{d,u}^2)$, so the mean and variance of y conditional on \hat{h} are

$$\mu_{y|\hat{h}} = 0 \quad (46)$$

$$\sigma_{y|\hat{h}}^2 = E_0 (|\hat{h}|^2 + \sigma_{d,u}^2) + \sigma_n^2 \quad (47)$$

$I(x; y|\hat{h})$ is maximized when $y|\hat{h}$ is Gaussian distributed. However, $y|\hat{h}$ is not Gaussian distributed in practise. In this case,

$$-E_y [\log p(y|\hat{h})] \leq \log (\pi e (E_0 (|\hat{h}|^2 + \sigma_{d,u}^2) + \sigma_n^2)) \quad (48)$$

Substituting (45) and (48) into (42) completes the proof.

D. Proof of Theorem 3

Since h is Rician distributed with parameters $\nu = \sqrt{\frac{K_R}{K_R+1}}$ and $\sigma = \sqrt{\frac{1}{2(K_R+1)}}$, i.e., $h \sim \text{Rice} \left(\sqrt{\frac{K_R}{K_R+1}}, \sqrt{\frac{1}{2(K_R+1)}} \right)$. Hence, $|\sqrt{2(K_R+1)}h|^2$ accords with a noncentral chi-squared distribution with two degrees of freedom and non-centrality parameter $2K_R$.

Combining (11), (20) and (22), it could be obtained that

$$\mathbf{R}_{\hat{d},u} = \frac{K_R}{K_R+1} \mathbf{I}_{N_p} + \mathbf{R}_{\hat{d},u}^{\text{DIF}} = \mathbf{I}_{N_p} - \mathbf{\Psi}_{ee,u} \quad (49)$$

From (49), $E[|\hat{h}|^2]$ is given by

$$E[|\hat{h}|^2] = \lim_{N_p \rightarrow \infty} \frac{1}{N_p} \mathbf{R}_{\hat{d},u} = 1 - \sigma_{d,u}^2 \quad (50)$$

$|h|^2$ and $|\hat{h}|^2$ are of highly correlated distributions. Moreover, (50) implies that $E[|\hat{h}|^2] \approx E[|h|^2]$. As a consequence, it is practical to use $|h|^2$ to approximate $|\hat{h}|^2$. Therefore,

$$C(\hat{h}, \gamma)_u \approx \log \left(1 + \frac{|\sqrt{2(K_R+1)}h|^2}{2(K_R+1)(\sigma_{d,u}^2 + \frac{1}{\gamma})} \right) \quad (51)$$

and

$$\eta \approx (1 - \delta) \int_0^{+\infty} \log \left(1 + \frac{x}{2(K_R+1)(\sigma_{d,u}^2 + \frac{1}{\gamma})} \right) p(x) dx \quad (52)$$

where $p(x)$ is the probability density function of $|\sqrt{2(K_R+1)}h|^2$, which could be expressed as

$$p(x) = \frac{1}{2} e^{-\frac{x+2K_R}{2}} I_0(\sqrt{2K_R x}) \quad (53)$$

$I_n(z)$ is n -th order modified Bessel function of the first kind.

REFERENCES

- [1] A. Abdi and M. Kaveh, "A space-time correlation model for multielement antenna systems in mobile fading channels," *Selected Areas in Communications, IEEE Journal on*, vol. 20, no. 3, pp. 550–560, 2002.
- [2] IST-WINNER, "Deliverable 1.1.2 v.1.2, WINNER II Channel Models," Tech. Rep., 2008 (<http://projects.celtic-initiative.org/winner+/deliverables.html>), Tech. Rep., 2007.
- [3] X. Ma, G. B. Giannakis, and S. Ohno, "Optimal training for block transmissions over doubly selective wireless fading channels," *Signal Processing, IEEE Transactions on*, vol. 51, no. 5, pp. 1351–1366, 2003.
- [4] R. Negi and J. Cioffi, "Pilot tone selection for channel estimation in a mobile OFDM system," *Consumer Electronics, IEEE Transactions on*, vol. 44, no. 3, pp. 1122–1128, 1998.
- [5] I. Barhumi, G. Leus, and M. Moonen, "Optimal training design for MIMO OFDM systems in mobile wireless channels," *Signal Processing, IEEE Transactions on*, vol. 51, no. 6, pp. 1615–1624, 2003.
- [6] N. Sun and J. Wu, "Maximizing spectral efficiency for high mobility systems with imperfect channel state information," *Wireless Communications, IEEE Transactions on*, 2014, accepted (to be published).
- [7] M. Grossglauser and D. Tse, "Mobility increases the capacity of ad-hoc wireless networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3. IEEE, 2001, pp. 1360–1369.
- [8] L. Sun and W. Wang, "On latency distribution and scaling: from finite to large cognitive radio networks under general mobility," in *INFOCOM, 2012 Proceedings IEEE*. IEEE, 2012, pp. 1287–1295.
- [9] X. Lu, J. L. Tiffany, Y. Liu, and C. Yuen, "A novel siso trellis strategy for relaying distorted signals in wireless networks," in *Communications (ICC), 2013 IEEE International Conference on*. IEEE, 2013, pp. 5247–5251.
- [10] H. Zhang, Z. Zhang, and H. Dai, "Gossip-based information spreading in mobile networks," *Wireless Communications, IEEE Transactions on*, vol. 12, no. 11, pp. 5918–5928, 2013.
- [11] L. Sun and W. Wang, "On distribution and limits of information dissemination latency and speed in mobile cognitive radio networks," in *INFOCOM, 2011 Proceedings IEEE*. IEEE, 2011, pp. 246–250.
- [12] X. Chen and C. Yuen, "Efficient resource allocation in a rateless-coded mu-mimo cognitive radio network with qos provisioning and limited feedback," *Vehicular Technology, IEEE Transactions on*, vol. 62, no. 1, pp. 395–399, 2013.
- [13] X. Chen and H. Chen, "Interference-aware resource control in multi-antenna cognitive ad hoc networks with heterogeneous delay constraints," *IEEE Communication Letter*, vol. 17, no. 6, pp. 1184–1187, 2013.
- [14] D. Gesbert, M. Shafi, D.-s. Shiu, P. J. Smith, and A. Naguib, "From theory to practice: an overview of MIMO space-time coded wireless systems," *Selected Areas in Communications, IEEE Journal on*, vol. 21, no. 3, pp. 281–302, 2003.
- [15] A. F. Molisch, "A generic model for MIMO wireless propagation channels in macro-and microcells," *Signal Processing, IEEE Transactions on*, vol. 52, no. 1, pp. 61–71, 2004.
- [16] K. I. Pedersen, J. B. Andersen, J. P. Kermoal, and P. Mogensen, "A stochastic multiple-input-multiple-output radio channel model for evaluation of space-time coding algorithms," in *Vehicular Technology Conference, 2000. IEEE-VTS Fall VTC 2000. 52nd*, vol. 2. IEEE, 2000, pp. 893–897.
- [17] J. Wu and N. Sun, "Optimum sensor density in distortion-tolerant wireless sensor networks," *Wireless Communications, IEEE Transactions on*, vol. 11, no. 6, pp. 2056–2064, 2012.
- [18] W. M. Gifford, M. Z. Win, and M. Chiani, "Diversity with practical channel estimation," *Wireless Communications, IEEE Transactions on*, vol. 4, no. 4, pp. 1935–1947, 2005.
- [19] J. Wu and C. Xiao, "Optimal diversity combining based on linear estimation of Rician fading channels," *Communications, IEEE Transactions on*, vol. 56, no. 10, pp. 1612–1615, 2008.
- [20] H. Gazzah, P. A. Regalia, and J.-P. Delmas, "Asymptotic eigenvalue distribution of block toeplitz matrices and application to blind simo channel identification," *Information Theory, IEEE Transactions on*, vol. 47, no. 3, pp. 1243–1251, 2001.